

Protein Sequence Searches - February 2005

All of the sequence databases on ABSS have recently been updated.

- Please note that the curators of the UniProt database have purged some temporary accession numbers from the most recent version of UniProt. These sequences have been assigned new permanent accession numbers. The new UniProt record may not contain the previous temporary accession number.
- If you encounter an accession number from an older search run against UniProt (results file extension **.rup**) that can no longer be found in the database, the permanent record with the new accession number can be found by searching the old accession number in the UniProt Protein Archive database (UniPARC) at:

<http://www.pir.uniprot.org/database/archive.shtml>

If you have any questions regarding this information or your results, please contact any STIC searcher.

When submitting sequence search results for scanning into IFW, please include a copy of this attachment to assist any future Examiners or members of the public who may encounter UniProt temporary accession numbers.

This Page Blank (uspto)

November 2005

Published_Applications Nucleic Acid and Published_Applications Amino Acid database searches now generate two sets of results each. The Published_Applications databases have been split into two parts to reduce the amount of time required for their daily updates. This results in more machine time being available for processing searches.

Newly published applications will appear in the Published_Applications_New databases: older published applications make up the Published_Applications_Main databases.

Searches run against Nucleic Acid Published_Applications produce two sets of results, with the extensions **.rnpbm** (Published_Applications_NA_Main) and **.rnpbn** (Published_Applications_NA_New). Searches run against Amino Acid Published_Applications produce two sets of results, with the extensions **.rapbm** (Published_Applications_AA_Main) and **.rapbn** (Published_Applications_AA_New).

This Page Blank (up to)

GenCore version 5.1.7
 Copyright (c) 1993 - 2006 Bioceleration Ltd.
 OM protein - protein search, using sw model
 Run on: March 2, 2006, 18:36:26 ; Search time 123.608 Seconds
 (without alignments)
 2100.780 Million cell updates/sec
 Title: US-10-693-367-2
 Perfect score: 3090
 Sequence: 1 MGKPKRVEISAPNSFEHR.....LAKGPPASIVPMQNRTR 591
 Scoring table: BLOSUM62
 Gapop 10.0 , Gapext 0.5
 Searched: 2443163 seqs, 439378781 residues
 Total number of hits satisfying chosen parameters: 2443163
 Minimum DB seq length: 0
 Maximum DB seq length: 2000000000
 Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing first 45 summaries

Database : A_Geneseq_21.*
 1: Geneseq1980s:*
 2: Geneseq1990s:*
 3: Geneseq2000s:*
 4: Geneseq2001s:*
 5: Geneseq2002s:*
 6: Geneseq2003as:*
 7: Geneseq2003bs:*
 8: Geneseq2004s:*
 9: Geneseq2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | DB ID | Description |
|------------|--------|---------------|--------|-------|--------------------|
| 1 | 3090 | 100.0 | 591 | 2 | AAY55964 Full leng |
| 2 | 3090 | 100.0 | 591 | 3 | AAY59128 Human ser |
| 3 | 3090 | 100.0 | 591 | 7 | ADC37307 Nuclear f |
| 4 | 3090 | 100.0 | 591 | 7 | ADP89973 Human can |
| 5 | 3090 | 100.0 | 591 | 7 | ADP45079 Human kin |
| 6 | 3090 | 100.0 | 591 | 8 | ADH23362 Human p21 |
| 7 | 3090 | 100.0 | 591 | 8 | ADH2101 Tumour-as |
| 8 | 3090 | 100.0 | 591 | 9 | ADW71912 Human p21 |
| 9 | 3090 | 100.0 | 591 | 9 | ADW06144 Cyclin-de |
| 10 | 3090 | 100.0 | 591 | 9 | ADZ64110 Human can |
| 11 | 3090 | 100.0 | 591 | 9 | ADG19308 Novel hum |
| 12 | 3075 | 99.5 | 588 | 4 | ADH23359 Human p21 |
| 13 | 2860 | 92.6 | 593 | 8 | ADJ96654 Human ste |
| 14 | 2674.5 | 86.6 | 522 | 8 | ABM84642 Human dia |
| 15 | 2674.5 | 86.6 | 522 | 8 | ABM84644 Human dia |
| 16 | 2674.5 | 86.6 | 522 | 8 | ABM84645 Human dia |
| 17 | 2662.5 | 86.2 | 516 | 9 | ADZ64115 Human can |
| 18 | 2662 | 86.1 | 517 | 8 | ABM84648 Human dia |
| 19 | 2662 | 86.1 | 517 | 8 | ABM84647 Human dia |
| 20 | 2662 | 86.1 | 517 | 8 | ABM84643 Human dia |
| 21 | 2662 | 86.1 | 517 | 8 | ABM84646 Human dia |
| 22 | 2558 | 82.8 | 501 | 7 | ADC37309 Nuclear f |
| 23 | 2358 | 82.8 | 501 | 9 | ADZ64111 Human can |
| 24 | 2202.5 | 71.3 | 438 | 7 | ADC37305 Nuclear f |

| | | | | | |
|----|--------|------|-----|---|--------------------|
| 25 | 2202.5 | 71.3 | 438 | 9 | ADZ64112 Human can |
| 26 | 2194.5 | 71.0 | 438 | 8 | ADR39763 Human kin |
| 27 | 2113.5 | 68.4 | 426 | 9 | ADZ64114 Human can |
| 28 | 2073 | 67.1 | 398 | 2 | AAY55941 Human PAK |
| 29 | 2069 | 67.0 | 405 | 9 | ADZ64113 Human can |
| 30 | 1748 | 56.6 | 719 | 7 | ADJ83007 Murine ma |
| 31 | 1739 | 56.3 | 719 | 7 | ADJ83008 Human mam |
| 32 | 1734 | 56.1 | 719 | 4 | AAM38963 Human pol |
| 33 | 1731 | 56.0 | 719 | 4 | AAB65705 Novel pro |
| 34 | 1731 | 56.0 | 719 | 7 | ADC37451 Nuclear f |
| 35 | 1731 | 56.0 | 719 | 7 | ADF45080 Human kin |
| 36 | 1731 | 56.0 | 719 | 8 | ADI29314 Human P21 |
| 37 | 1730 | 56.0 | 719 | 8 | AAE02187 Human p21 |
| 38 | 1549 | 50.1 | 457 | 8 | ADH42207 Novel hum |
| 39 | 1549 | 50.1 | 457 | 8 | ADH42215 Novel hum |
| 40 | 1549 | 50.1 | 457 | 8 | ADH42217 Novel hum |
| 41 | 1545 | 50.0 | 457 | 8 | ADH42213 Novel hum |
| 42 | 1535 | 49.7 | 457 | 8 | ADH42211 Novel hum |
| 43 | 1502 | 48.6 | 293 | 7 | ADE15851 PAK4KD pr |
| 44 | 1497.5 | 48.5 | 639 | 8 | ADH23364 Fruit fly |
| 45 | 1494.5 | 48.4 | 635 | 8 | ADH23360 Fruit fly |

ALIGNMENTS

RESULT 1
 AAY55964
 ID AAY55964 standard; protein; 591 AA.
 XX
 AC AAY55964;
 XX
 DT 18-FEB-2000 (first entry)
 XX
 DE Full length human PAK5 protein.
 XX
 KW Antirheumatic; antiarthritic; antiinflammatory; antiallergic; osteopathic;
 KW antipsoriatic; antiarteriosclerotic; antidiabetic; antidiabetic;
 KW neuroprotective; cardiant; cerebroprotective; cytostatic; antidiabetic;
 KW vulnery; SPE20; protein kinase; STLK3; STLK4; STLK5; STLK6; STLK7;
 KW ZC1; ZC2; ZC3; ZC4; KHS2; SULUI; SULU3; GEK2; PAK4; PAK5; antagonist;
 KW antibody; gene therapy; rheumatoid arthritis; artherosclerosis; asthma;
 KW inflammatory bowel disease; Crohn's disease; osteoarthritis; psoriasis;
 KW rhinitis; autoimmunity; organ transplantation; multiple sclerosis;
 KW myocardial infarction; cardiovascular disease; stroke; renal failure;
 KW oxidative stress-related neurodegenerative disorder; Parkinson's disease;
 KW amyotrophic lateral sclerosis; Leigh syndrome; cancer; cardiomyopathy;
 KW ischemic disorder; inflammation; diabetes mellitus; fibrosis; mitosis;
 KW mesangial disorder; growth regulation; wound healing; T cell activation;
 KW immunosuppressant.
 KW Homo sapiens.
 OS
 XX
 XX WO9953036-A2.
 XX
 XX 21-OCT-1999.
 XX
 XX 13-APR-1999; 99WO-US008150.
 XX
 PR 14-APR-1998; 98US-0081784P.
 XX
 XX (SUGC-) SUGEN INC.
 XX
 XX Plowman G, Martinez R, Whyte D;
 XX WPI; 1999-611301/52.
 XX N-PSDB; AA240538.
 XX
 XX Novel kinase-related polypeptides used for the diagnosis and treatment of
 XX kinase-related diseases and disorders.
 XX Claim 11; Page 366-368; 387pp; English.

CC This sequence represents a novel STE20-related protein kinase. The
 CC invention relates to nucleic acid molecule encoding a kinase polypeptide
 CC selected from STUK2, STUK3, STUK4, STUK5, STUK6, STUK7, ZC1, ZC2, ZC3,
 CC ZC4, KHS2, SULJ1, SULJ3, GEK2, PAK4 and PAK5. The proteins are used to
 CC identify agonists and antagonists, and to raise antibodies. The
 CC polynucleotides are useful in gene therapy protocols. The polynucleotides,
 CC polypeptides, antibodies, antagonists and agonists may be used to treat
 CC diseases such as immune-related disorders and diseases (e.g. rheumatoid
 CC arthritis, atherosclerosis, chronic inflammatory bowel disease (e.g.
 CC Crohn's disease), asthma, osteoarthritis, psoriasis, atherosclerosis,
 CC rhinitis, autoimmunity, and organ transplantation, chronic inflammatory
 CC pelvic disease, multiple sclerosis, organ transplantation, myocardial
 CC infarction, cardiovascular disease, stroke, renal failure, oxidative
 CC stress-related neurodegenerative disorders (e.g. amyotrophic lateral
 CC sclerosis, Parkinson's disease and Leigh syndrome), cancer,
 CC cardiomyopathies, ischemic disorders, inflammatory disorders, diabetes
 CC mellitus, fibrotic and mesangial disorders. The proteins may also be
 CC useful for cell growth regulation (e.g. in wound healing), T cell
 CC activation, mitosis control, and as immunosuppressants
 XX
 XX Sequence 591 AA;

Query Match 100.0%; Score 3090; DB 2; Length 591;
 Best Local Similarity 100.0%; Pred. No. 3.2e-216;
 Matches 591; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MFGKKRKRVEISAPNFEHRVHTGFDQHEQKFTGLPROMQSLIEESARRPKPLVDPAIT 60
 DB 1 MFGKKRKRVEISAPNFEHRVHTGFDQHEQKFTGLPROMQSLIEESARRPKPLVDPAIT 60

QY 61 SIQGPAPKTIIVRGSKAKDGALTLLDEFENSVTRNSLRDSDPPPPARARQENGMPEE 120
 DB 61 SIQGPAPKTIIVRGSKAKDGALTLLDEFENSVTRNSLRDSDPPPPARARQENGMPEE 120

QY 121 PATTARGGPKAGSRGRFAGHSEAGGGSDRRRAGPEKPKSRSGSGPQSSRDKRPL 180
 DB 121 PATTARGGPKAGSRGRFAGHSEAGGGSDRRRAGPEKPKSRSGSGPQSSRDKRPL 180

QY 181 SGPDVGTTPQAGLASKAKLAAGRPNTYPRADTDHPSRGAQGEHDVAPNGPSAGGLAIP 240
 DB 181 SGPDVGTTPQAGLASKAKLAAGRPNTYPRADTDHPSRGAQGEHDVAPNGPSAGGLAIP 240

QY 241 QSSSSSSRPPTTRARGAPSGVLGPHASEPQLAPPACTPAAPVPGPPGPRSPQRPQVS 300
 DB 241 QSSSSSSRPPTTRARGAPSGVLGPHASEPQLAPPACTPAAPVPGPPGPRSPQRPQVS 300

QY 301 HEQFRAALQLVVDGDPDRSYLDNFIKIGSGTGIVCIAIVRSSGKLVAVKMDLRKQRR 360
 DB 301 HEQFRAALQLVVDGDPDRSYLDNFIKIGSGTGIVCIAIVRSSGKLVAVKMDLRKQRR 360

QY 361 ELLFNEVIMRDYQHENVVEMYSYLVGDDELVWMEFLEGGALTIDIVTHTRNNEBQIAAV 420
 DB 361 ELLFNEVIMRDYQHENVVEMYSYLVGDDELVWMEFLEGGALTIDIVTHTRNNEBQIAAV 420

QY 421 CLAVLQALSVLHAQGVIIHRDIKSDSILLTHDGRVKLSDFGCAQVSKVPRKSLVGTTPY 480
 DB 421 CLAVLQALSVLHAQGVIIHRDIKSDSILLTHDGRVKLSDFGCAQVSKVPRKSLVGTTPY 480

QY 481 WNAPELISRLPYGPEVDIWSLGIWMTVMVDGPPVFNPEPPLKAMKMRDNLPLRLKHLK 540
 DB 481 WNAPELISRLPYGPEVDIWSLGIWMTVMVDGPPVFNPEPPLKAMKMRDNLPLRLKHLK 540

QY 541 VSPSLKGFLLRLVDRPAQATAAELLKHPFLAKAGPPASIVPLMRQNRTR 591
 DB 541 VSPSLKGFLLRLVDRPAQATAAELLKHPFLAKAGPPASIVPLMRQNRTR 591

RESULT 2
 AAY59128
 ID AAY59128 standard; protein; 591 AA.
 XX
 AC AAY59128;
 XX

DT 08-MAR-2000 (first entry)
 XX Human serine/threonine kinase, PAK4.
 DE PAK4; serine/threonine kinase; GTPase; intracellular signal cascade; Rac;
 XX Cdc42H; morphogenesis; mitogenesis; JNK; p38 MAP kinase; human;
 KW actin polymerization; filopodia; cancer; arthritis.
 KW
 XX Homo sapiens.
 OS
 XX WO9963073-A1.
 PN
 XX 09-DEC-1999.
 PD
 XX 21-MAY-1999; 99WO-US011341.
 PF
 XX 21-MAY-1998; 98US-00082737.
 PR
 XX (UYCO) UNIV COLUMBIA NEW YORK.
 PA
 XX Minden A;
 PI
 XX WPI; 2000-072881/06.
 DR
 XX N-PSDB; AAZ40657.
 DX
 XX Novel mammalian nucleic acid useful for treating cancer and arthritis.
 PT
 XX Claim 8; Fig 1A-B; 95pp; English.
 PS
 XX
 XX The invention relates to an isolated mammalian nucleic acid that encodes
 CC PAK4, a novel serine/threonine kinase or its mutant homolog. PAK4 is an
 CC effector for the GTPases Rac and Cdc42Hs which are involved in
 CC intracellular signal cascades, morphogenesis and mitogenesis, and
 CC activate the JNK and p38 MAP kinase pathways. Inhibiting interaction of
 CC PAK4 with these enzymes will thus result in inhibition of actin
 CC polymerization and formation of filopodia. The PAK4 nucleic acid used for
 CC recombinant production of the protein, and as a source of probes for
 CC identifying homologous sequences and of (antisense oligonucleotides for
 CC inhibiting PAK4 expression. The protein, or its fragments, are used to
 CC raise specific antibodies and these are useful as ligands for therapeutic
 CC inhibition of interaction between PAK4 and its native binding partners.
 CC Inhibition of PAK4 activity or expression is used for treatment of cancer
 CC and arthritis. The present sequence represents the human serine/threonine
 CC kinase, PAK4
 CC
 XX Sequence 591 AA;

Query Match 100.0%; Score 3090; DB 3; Length 591;
 Best Local Similarity 100.0%; Pred. No. 3.2e-216;
 Matches 591; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MFGKKRKRVEISAPNFEHRVHTGFDQHEQKFTGLPROMQSLIEESARRPKPLVDPAIT 60
 DB 1 MFGKKRKRVEISAPNFEHRVHTGFDQHEQKFTGLPROMQSLIEESARRPKPLVDPAIT 60

QY 61 SIQGPAPKTIIVRGSKAKDGALTLLDEFENSVTRNSLRDSDPPPPARARQENGMPEE 120
 DB 61 SIQGPAPKTIIVRGSKAKDGALTLLDEFENSVTRNSLRDSDPPPPARARQENGMPEE 120

QY 121 PATTARGGPKAGSRGRFAGHSEAGGGSDRRRAGPEKPKSRSGSGPQSSRDKRPL 180
 DB 121 PATTARGGPKAGSRGRFAGHSEAGGGSDRRRAGPEKPKSRSGSGPQSSRDKRPL 180

QY 181 SGPDVGTTPQAGLASKAKLAAGRPNTYPRADTDHPSRGAQGEHDVAPNGPSAGGLAIP 240
 DB 181 SGPDVGTTPQAGLASKAKLAAGRPNTYPRADTDHPSRGAQGEHDVAPNGPSAGGLAIP 240

QY 241 QSSSSSSRPPTTRARGAPSGVLGPHASEPQLAPPACTPAAPVPGPPGPRSPQRPQVS 300
 DB 241 QSSSSSSRPPTTRARGAPSGVLGPHASEPQLAPPACTPAAPVPGPPGPRSPQRPQVS 300

QY 301 HEQFRAALQLVVDGDPDRSYLDNFIKIGSGTGIVCIAIVRSSGKLVAVKMDLRKQRR 360
 DB 301 HEQFRAALQLVVDGDPDRSYLDNFIKIGSGTGIVCIAIVRSSGKLVAVKMDLRKQRR 360

| Result No. | Query | | | DB | ID | Description |
|------------|--------|-------|--------|----|--------|---------------------|
| | Score | Match | Length | | | |
| 1 | 1096.5 | 35.5 | 540 | 2 | T19956 | hypothetical prote |
| 2 | 1090.5 | 35.3 | 542 | 2 | T19852 | hypothetical prote |
| 3 | 917 | 29.7 | 544 | 2 | AS7597 | beta-p21-activated |
| 4 | 907 | 29.4 | 544 | 2 | T49376 | p21 activated kina |
| 5 | 895.5 | 29.0 | 525 | 2 | S58682 | protein kinase, p2 |
| 6 | 892.5 | 28.9 | 545 | 2 | G01773 | p21-activated prot |
| 7 | 892 | 28.9 | 544 | 2 | S40482 | serine/threonine-s |
| 8 | 846.5 | 27.4 | 1230 | 2 | T18256 | probable serine/th |
| 9 | 846.5 | 27.4 | 1230 | 2 | T18259 | serine/threonine p |
| 10 | 837 | 27.1 | 939 | 2 | S28394 | probable serine/th |
| 11 | 823.5 | 26.7 | 658 | 2 | T39500 | serine/threonine-s |
| 12 | 814.5 | 26.4 | 658 | 2 | S60170 | protein kinase Pak |
| 13 | 811.5 | 26.3 | 622 | 2 | T15467 | hypothetical prote |
| 14 | 760 | 24.6 | 842 | 2 | S60402 | protein kinase ClaA |
| 15 | 722 | 23.4 | 378 | 2 | T26684 | hypothetical prote |
| 16 | 697 | 22.6 | 589 | 2 | T38086 | serine/threonine-p |
| 17 | 651.5 | 21.1 | 655 | 2 | S51884 | probable protein k |
| 18 | 505 | 16.3 | 836 | 2 | B95716 | probable serine/th |
| 19 | 498 | 16.1 | 471 | 2 | T39232 | probable serine th |
| 20 | 494 | 16.0 | 1102 | 2 | JC6316 | probable protein k |
| 21 | 478 | 15.5 | 653 | 2 | T34356 | hypothetical prote |
| 22 | 476 | 15.4 | 819 | 2 | AS3714 | protein kinase (EC |
| 23 | 470.5 | 15.2 | 829 | 2 | T29372 | hypothetical prote |
| 24 | 464 | 15.0 | 426 | 2 | S71886 | Ste20-like protein |
| 25 | 463 | 15.0 | 1231 | 2 | T18532 | serine/threonine pr |
| 26 | 462 | 15.0 | 1233 | 2 | T14157 | serine/threonine p |
| 27 | 460 | 14.9 | 1206 | 2 | T34021 | protein kinase SK2 |
| 28 | 458.5 | 14.7 | 690 | 2 | C96572 | protein-F12M16.4 [|
| 29 | 455.5 | 14.7 | 1080 | 2 | S48944 | hypothetical prote |

[illegible]

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 2, 2006, 18:36:41 ; Search time 137.127 Seconds
(without alignments)
3040.728 Million cell updates/sec

Title: US-10-693-367-2

Perfect score: 3090

Sequence: 1 MFGRKKRVRISAPNSFEHR.....LAKAGFPASIVPLMRQNRTR 591

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt 05.80.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|--------|-------------|--------|----------------|---------------------|
| 1 | 3090 | 100.0 | 591 | 1 PAK4_HUMAN | O96013 homo sapien |
| 2 | 2860 | 92.6 | 593 | 1 PAK4_MOUSE | Q8btw9 mus musculus |
| 3 | 2860 | 92.6 | 597 | 2 Q62PX0_MOUSE | Q62px0 mus musculus |
| 4 | 2853 | 92.3 | 593 | 2 Q80Z97_MOUSE | Q80z97 mus musculus |
| 5 | 2558 | 82.8 | 501 | 2 Q8N4E1_HUMAN | Q8n4e1 homo sapien |
| 6 | 2202.5 | 71.3 | 438 | 2 Q8NDE3_HUMAN | Q8nde3 homo sapien |
| 7 | 2202.5 | 71.3 | 467 | 2 Q9ULS8_HUMAN | Q9uls8 homo sapien |
| 8 | 2126.5 | 68.8 | 650 | 2 Q5PQ75_XENLA | Q5pq75 xenopus lae |
| 9 | 2124.5 | 68.8 | 650 | 2 Q90W62_XENLA | Q90w62 xenopus lae |
| 10 | 2031 | 65.7 | 663 | 2 Q6GM11_BRARE | Q6gm11 brachydanio |
| 11 | 1967 | 63.7 | 407 | 2 Q8K0U2_MOUSE | Q8k0u2 mus musculus |
| 12 | 1949 | 63.1 | 589 | 2 Q4STA5_TETNG | Q4sta5 tetraodon n |
| 13 | 1748 | 56.6 | 719 | 2 Q8C015_MOUSE | Q8c015 mus musculus |
| 14 | 1745 | 56.5 | 711 | 2 Q6NVJ4_BRARE | Q6nvj4 brachydanio |
| 15 | 1740 | 56.3 | 719 | 2 Q5R7U3_PONGY | Q5r7u3 pongo pygma |
| 16 | 1740 | 56.3 | 719 | 2 Q8BVB0_MOUSE | Q8bvb0 mus musculus |
| 17 | 1731 | 56.0 | 719 | 1 PAK7_HUMAN | Q9p286 homo sapien |
| 18 | 1731 | 56.0 | 719 | 2 Q8TB93_HUMAN | Q8tb93 homo sapien |
| 19 | 1731 | 56.0 | 719 | 2 Q5W115_HUMAN | Q5w115 homo sapien |
| 20 | 1730 | 56.0 | 719 | 2 Q6RWS7_MOUSE | Q6rws7 mus musculus |
| 21 | 1672 | 54.1 | 908 | 2 Q4SCY3_TETNG | Q4scy3 tetraodon n |
| 22 | 1507 | 48.8 | 616 | 2 Q7QGS5_ANOGA | Q7qgs5 anopheles g |
| 23 | 1501.5 | 48.6 | 639 | 1 PAKM_DROME | Q9vxe5 drosophila |
| 24 | 1476.5 | 47.8 | 681 | 2 Q5R8Z4_PONGY | Q5r8z4 pongo pygma |
| 25 | 1473.5 | 47.7 | 681 | 1 PAK6_HUMAN | Q9nqu5 homo sapien |
| 26 | 1405 | 45.5 | 625 | 2 Q5SQE8_BRARE | Q5sqe8 brachydanio |
| 27 | 1241 | 40.2 | 240 | 2 Q8NCH5_HUMAN | Q8nch5 homo sapien |
| 28 | 1160 | 37.5 | 229 | 2 Q9CS71_MOUSE | Q9cs71 mus musculus |
| 29 | 1096.5 | 35.5 | 540 | 2 Q9U3M1_CABEL | Q9u3m1 caenorhabdi |
| 30 | 1096 | 35.5 | 286 | 2 Q4R3T4_MACFA | Q4r3t4 macaca fasc |
| 31 | 1090.5 | 35.3 | 542 | 2 Q18637_CABEL | Q18637 caenorhabdi |

| | | | | | | |
|----|-------|------|-----|---|--------------|---------------------|
| 32 | 1005 | 32.5 | 499 | 2 | Q60MNO_CABER | Q60mno caenorhabdi |
| 33 | 950 | 30.7 | 662 | 2 | Q7QD76_ANOGA | Q7qd76 anopheles g |
| 34 | 937.5 | 30.3 | 704 | 2 | Q24190_DROME | Q24190 drosophila |
| 35 | 937.5 | 30.3 | 704 | 2 | Q9V113_DROME | Q9v113 drosophila |
| 36 | 936 | 30.3 | 827 | 2 | Q61WN3_USTMA | Q61wn3 ustilago ma |
| 37 | 936 | 30.3 | 827 | 2 | Q4PBV7_USTMA | Q4pbv7 ustilago ma |
| 38 | 932.5 | 30.2 | 704 | 2 | Q24213_DROME | Q24213 drosophila |
| 39 | 923 | 29.9 | 848 | 2 | Q5BBL3_EMENI | Q5bbl3 aspergillus |
| 40 | 917 | 29.7 | 544 | 1 | PAK3_RAT | Q62829 rattus norv |
| 41 | 907 | 29.4 | 577 | 2 | Q803Z0_BRARE | Q803z0 brachydanio |
| 42 | 907 | 29.4 | 986 | 1 | Q4WHP3_ASPFU | Q4whp3 aspergillus |
| 43 | 906.5 | 29.3 | 559 | 1 | PAK3_MOUSE | Q61036 mus musculus |
| 44 | 905 | 29.3 | 577 | 2 | Q6P0I7_BRARE | Q6p0i7 brachydanio |
| 45 | 904 | 29.3 | 524 | 1 | PAK2_HUMAN | Q13177 homo sapien |

ALIGNMENTS

RESULT 1

ID PAK4_HUMAN STANDARD; PRT; 591 AA.
AC O96013; Q9BU33;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 13-SEP-2005 (Rel. 48, Last annotation update)

DE Serine/threonine-protein kinase PAK 4 (EC 2.7.1.37) (p21-activated

kinase 4) (PAK-4).

GN Name=PAK4;

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;

OC Homo.

OX NCBI_TaxID=9606;

RN [1]_

RP NUCLEOTIDE SEQUENCE (ISOFORM 1).

RX MEDLINE=99043860; PubMed=9822598; DOI=10.1093/emboj/17.22.6527;

RA Abo A., Qu J., Cammarano M.S., Dan C., Fritsch A., Baud V.,

RA Belisle B., Minden A.;

RT "PAK4, a novel effector for Cdc42Hs, is implicated in the

reorganization of the actin cytoskeleton and in the formation of

filopodia.";

RT EMBO J. 17:6527-6540 (1998).

RN [2]

RP NUCLEOTIDE SEQUENCE (ISOFORM 1).

RA Submitted (MAY-1997) to the EMBL/GenBank/DBSJ databases.

RN [3]

RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORMS 1 AND 2).

RC TISSUE=Eye, Pancreas, and Placenta;

RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,

RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Prange C.,

RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

RA Villalón D.J., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

RA Blakesley R.W., Trinchman J.W., Green E.D., Dickson M.C.,

RA Rodriguez A.C., Grinwood J., Schmutz J., Myers R.M.,

RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,

RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;

RT "Generation and initial analysis of more than 15,000 full-length human

and mouse cDNA sequences."

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).

CC -!- FUNCTION: Activates the JNK pathway. Implicated in the

reorganization of the actin cytoskeleton and in the formation of

CC filopodia.
CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -!- SUBUNIT: Interacts tightly with GTP-bound but not GDP-bound
CC CDC42/p21 and weakly with RAC1.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=O96013-1; Sequence=Displayed;
CC Name=2;
CC IsoId=O96013-2; Sequence=VSP_004892; VSP_004893;
CC Note=No experimental confirmation available;
CC -!- TISSUE SPECIFICITY: Highest expression in prostate, testis and
CC colon.
CC -!- PTM: Autophosphorylated when activated by CDC42/p21.
CC -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family. STE20
CC subfamily.
CC -!- SIMILARITY: Contains 1 CRIB domain.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
CC EMBL; AJ011855; CAA09820.1; -; mRNA.
CC EMBL; AF005046; AAD01210.1; -; mRNA.
CC EMBL; BC002921; AAH02921.1; -; mRNA.
CC EMBL; BC011368; AAH11368.1; -; mRNA.
CC EMBL; BC025282; AAH25282.1; -; mRNA.
CC HSSP; Q13153; 1F3M.
CC Ensembl; ENSG00000130669; Homo sapiens.
CC HGNC; HGNC:16059; PAK4.
CC H-InvDB; HIX0015110; -.
CC MIM; 605451; -.
CC GO; GO:0005794; C:Golgi apparatus; TAS.
CC GO; GO:0004672; F:protein kinase activity; NAS.
CC GO; GO:0006928; P:cell motility; TAS.
CC GO; GO:0007185; P:signal transduction; TAS.
CC InterPro; IPR000095; PAK box Rho bd.
CC InterPro; IPR000719; Prot Kinase.
CC InterPro; IPR008271; Ser thr_pkin AS.
CC InterPro; IPR002290; Ser_thr_pkinase.
CC Pfam; PF00786; PBD; 1.
CC Pfam; PF00069; Pkinase; 1.
CC ProDom; PD000001; Prot_kinase; 1.
CC SMART; SM00285; PBD; 1.
CC SMART; SM00220; S_TKc; 1.
CC PROSITE; PS0108; CRIB; 1.
CC PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE; PS00107; PROTEIN_KINASE_DOM; 1.
CC PROSITE; PS00108; PROTEIN_KINASE_ST; FALSE_NEG.
CC Alternative splicing; ATP-binding; Kinase; Nucleotide-binding;
CC phosphorylation; Serine/threonine-protein kinase; Transferase.
CC DOMAIN 11 24 CRIB
CC FT DOMAIN 321 572 Protein kinase.
CC FT NP_BIND 327 335 ATP (By similarity).
CC FT REGION 25 320 Linker.
CC FT ACT_SITE 440 450 Proton acceptor (By similarity).
CC FT BINDING 350 350 ATP (By similarity).
CC FT VARSP_LIC 120 120 E -> K (in isoform 2).
CC FT VARSP_LIC 121 285 /FTId=VSP_004892.
CC FT VARSP_LIC 121 285 Missing (in isoform 2).
CC FT VARSP_LIC 121 285 /FTId=VSP_004893.
CC SEQUENCE 591 AA; 64072 MW; 04C2A5C0B06427D5 CRC64;
CC
CC Query Match 100.08; Score 3090; DB 1; Length 591;
CC Best Local Similarity 100.08; Pred. No. 6.1e-127; Indels 0;
CC Matches 591; Conservative 0; Mismatches 0; Gaps 0;
CC
CC 1 MFGRKKRVEISAPSNFHRVHTGFDQHEQKFTGLPRWQSLIEESARRPKPLVDPAICIT 60
CC
CC 1 MFGRKKRVEISAPSNFHRVHTGFDQHEQKFTGLPRWQSLIEESARRPKPLVDPAICIT 60
CC

QY 61 SIQGPACTIVRGSGKADGALTLLDEFENMSVTRNSLRDSDPPPPARARQENGPPE 120
DB |||||
DB 61 SIQGPACTIVRGSGKADGALTLLDEFENMSVTRNSLRDSDPPPPARARQENGPPE 120
QY 121 PATTARGGPGKAGSRGPFAGHSEAGGSGDRRRAGPEKPKSSREGSGGPGPSSSRDKRPL 180
DB |||||
DB 121 PATTARGGPGKAGSRGPFAGHSEAGGSGDRRRAGPEKPKSSREGSGGPGPSSSRDKRPL 180
QY 181 SGPVGTPTQAGLAGAKLAAGRPFNTYPRADTTHPSRGAQGEHVDVAPNPGSAGGLAIP 240
DB |||||
DB 181 SGPVGTPTQAGLAGAKLAAGRPFNTYPRADTTHPSRGAQGEHVDVAPNPGSAGGLAIP 240
QY 241 QSSSSSRPPTTRARGAPSPGVLPAGHSEAGGSGDRRRAGPEKPKSSREGSGGPGPSSSRDKRPL 300
DB |||||
DB 241 QSSSSSRPPTTRARGAPSPGVLPAGHSEAGGSGDRRRAGPEKPKSSREGSGGPGPSSSRDKRPL 300
QY 301 HEQFRAALQLVDPGDRPSYLDNFIKIGEGSTGIVCTATVRSSGKLAVKMDLKKQRR 360
DB |||||
DB 301 HEQFRAALQLVDPGDRPSYLDNFIKIGEGSTGIVCTATVRSSGKLAVKMDLKKQRR 360
QY 361 ELLFNEVIMRDYQHENVVMYNSYLVDGLVWVMEFLEGGALTDIVTHTRMNEEQIAAV 420
DB |||||
DB 361 ELLFNEVIMRDYQHENVVMYNSYLVDGLVWVMEFLEGGALTDIVTHTRMNEEQIAAV 420
QY 421 CLAVLOALSVLHAQGVHRIKSDSILLTHDGRVKLSDFGCAQVSKVEPRKSLVGTTPY 480
DB |||||
DB 421 CLAVLOALSVLHAQGVHRIKSDSILLTHDGRVKLSDFGCAQVSKVEPRKSLVGTTPY 480
QY 481 WMAPELISRLPYGPEVDIWSLGINVIMVWDGEPYFNEPPLKAMKMTDNLPPRLKNLHK 540
DB |||||
DB 481 WMAPELISRLPYGPEVDIWSLGINVIMVWDGEPYFNEPPLKAMKMTDNLPPRLKNLHK 540
QY 541 VPSLKGFLDRLLVRDPAQRATAAELLKHPFLAKAGPPASIVPLMRQNRTR 591
DB |||||
DB 541 VPSLKGFLDRLLVRDPAQRATAAELLKHPFLAKAGPPASIVPLMRQNRTR 591

RESULT 2
PAK4 MOUSE
ID PAK4 MOUSE STANDARD; PRT; 593 AA.
AC Q8BTW9;
DT 05-JUL-2004 (Rel. 44, Created)
DT 05-JUL-2004 (Rel. 44, Last sequence update)
DT 13-SEP-2005 (Rel. 46, Last annotation update)
DE Serine/threonine-protein kinase PAK 4 (EC 2.7.1.37) (p21-activated
DE kinase 4) (PAK-4).
GN Name=PAK4;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1] NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RP STRAIN=NOD; TISSUE=Thymus;
RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA Nikaide I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA Blake J.A., Bradt D., Bruscia V., Chothia C., Corbani L.E., Cousins S.,
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,

| Result No. | Score | | Match | Length | DB | ID | Description |
|------------|--------|-------|-------|--------|---------------------|---------------------|-------------|
| | | | | | | | |
| 1 | 3090 | 100.0 | 591 | 2 | US-09-082-737-2 | Sequence 2, Appli | |
| 2 | 3090 | 100.0 | 591 | 2 | US-09-688-188B-103 | Sequence 103, Appli | |
| 3 | 3090 | 100.0 | 591 | 2 | US-09-718-032-2 | Sequence 2, Appli | |
| 4 | 3090 | 100.0 | 591 | 2 | US-09-291-417D-103 | Sequence 103, Appli | |
| 5 | 3090 | 100.0 | 591 | 2 | US-09-949-016-6665 | Sequence 6665, Ap | |
| 6 | 3090 | 100.0 | 620 | 2 | US-09-949-016-7206 | Sequence 7206, Ap | |
| 7 | 2073 | 67.1 | 398 | 2 | US-09-688-188B-30 | Sequence 30, Appli | |
| 8 | 2073 | 67.1 | 398 | 2 | US-09-291-417D-30 | Sequence 30, Appli | |
| 9 | 1473.5 | 47.7 | 681 | 2 | US-09-688-188B-29 | Sequence 29, Appli | |
| 10 | 1473.5 | 47.7 | 681 | 2 | US-09-291-417D-29 | Sequence 29, Appli | |
| 11 | 1473.5 | 47.7 | 681 | 2 | US-09-659-737A-10 | Sequence 10, Appli | |
| 12 | 1473.5 | 47.7 | 681 | 2 | US-10-885-921-10 | Sequence 10, Appli | |
| 13 | 1473.5 | 47.7 | 694 | 2 | US-09-949-016-10445 | Sequence 10445, A | |
| 14 | 1470.5 | 47.6 | 681 | 2 | US-09-765-815-2 | Sequence 2, Appli | |
| 15 | 1277 | 41.3 | 250 | 2 | US-09-082-737-3 | Sequence 3, Appli | |
| 16 | 1277 | 41.3 | 250 | 2 | US-09-718-032-3 | Sequence 3, Appli | |
| 17 | 1029 | 33.3 | 311 | 2 | US-09-659-737A-8 | Sequence 8, Appli | |
| 18 | 1029 | 33.3 | 311 | 2 | US-10-885-921-8 | Sequence 8, Appli | |
| 19 | 912 | 29.5 | 544 | 1 | US-08-935-760-2 | Sequence 2, Appli | |
| 20 | 909 | 29.4 | 544 | 2 | US-09-949-016-11562 | Sequence 11562, A | |
| 21 | 907 | 29.4 | 544 | 2 | US-09-688-188B-95 | Sequence 95, Appli | |
| 22 | 907 | 29.4 | 544 | 2 | US-09-291-417D-95 | Sequence 95, Appli | |
| 23 | 904 | 29.3 | 524 | 2 | US-09-538-092-1301 | Sequence 1301, Ap | |
| 24 | 900 | 29.1 | 524 | 1 | US-08-615-942A-2 | Sequence 2, Appli | |
| 25 | 900 | 29.1 | 524 | 2 | US-09-237-325-2 | Sequence 2, Appli | |
| 26 | 899 | 29.1 | 544 | 2 | US-08-559-397A-19 | Sequence 19, Appli | |
| 27 | 894 | 28.9 | 506 | 1 | US-08-369-780-2 | Sequence 2, Appli | |

Db 61 SIQCAPKTIVRGSKAKDGLTLLDDEFNNVTSRNSLRDSDPPPARAQENGMPEE 120
 QY 121 PATTARGGPKAGSRGRFAGHSEAGGSGDRRRRAGPEKPKSSREGSGGPQESSDKRPL 180
 Db 121 PATTARGGPKAGSRGRFAGHSEAGGSGDRRRRAGPEKPKSSREGSGGPQESSDKRPL 180
 QY 181 SGPDVGTTPQAGLASGAKLAAGRPNFTYPRADTDHPSRGAQCEPHDVAPNGPSAGGLAIP 240
 Db 181 SGPDVGTTPQAGLASGAKLAAGRPNFTYPRADTDHPSRGAQCEPHDVAPNGPSAGGLAIP 240
 QY 241 QSSSSSSRPPTRARGAPSGVLGPHASEPQLAPACTPAAPVPGPPGPRSPQRPQVRS 300
 Db 241 QSSSSSSRPPTRARGAPSGVLGPHASEPQLAPACTPAAPVPGPPGPRSPQRPQVRS 300
 QY 301 HEQFRAALQLVVDGPDPSYLDNFIKIGEGSTGIVCIATVRSSGKLVAVKMDLRKQRR 360
 Db 301 HEQFRAALQLVVDGPDPSYLDNFIKIGEGSTGIVCIATVRSSGKLVAVKMDLRKQRR 360
 QY 361 ELLFNEVIMRDYQHENVVEMYSYLVGDELWVMEFLEGGALTDIVTHTRMNEEQIAAV 420
 Db 361 ELLFNEVIMRDYQHENVVEMYSYLVGDELWVMEFLEGGALTDIVTHTRMNEEQIAAV 420
 QY 421 CLAVLQALSVLHAQGVTHRDIKSDSILLTHDGRVKLSDFGFCQAQVSKVPRKSLVGTPT 480
 Db 421 CLAVLQALSVLHAQGVTHRDIKSDSILLTHDGRVKLSDFGFCQAQVSKVPRKSLVGTPT 480
 QY 481 WMAPELISRLPYGPEVDIWSLGIWIMVMDGEPYFNEPPLKAMKMIKIRDLNPPRLKNLHK 540
 Db 481 WMAPELISRLPYGPEVDIWSLGIWIMVMDGEPYFNEPPLKAMKMIKIRDLNPPRLKNLHK 540
 QY 541 VSPSLKGFLLRDLVRDPAQRATAAELLKHPFLAKAGPPASIVPLMRQNRTR 591
 Db 541 VSPSLKGFLLRDLVRDPAQRATAAELLKHPFLAKAGPPASIVPLMRQNRTR 591

RESULT 2
 US-09-688-188B-103
 ; Sequence 103, Application US/09688188B
 ; Patent No. 6656716
 ; GENERAL INFORMATION:
 ; APPLICANT: PLOWMAN, GREGORY
 ; APPLICANT: MARTINEZ, RICARDO
 ; APPLICANT: WHYTE, DAVID
 ; TITLE OF INVENTION: STE20-RELATED PROTEIN KINASES
 ; FILE REFERENCE: 038602/0328
 ; CURRENT APPLICATION NUMBER: US/09/688,188B
 ; CURRENT FILING DATE: 2000-10-16
 ; PRIOR APPLICATION NUMBER: 09/291,417
 ; PRIOR FILING DATE: 1999-04-14
 ; PRIOR APPLICATION NUMBER: 60/081,784
 ; PRIOR FILING DATE: 1998-04-14
 ; NUMBER OF SEQ ID NOS: 155
 ; SOFTWARE: Patent in Ver. 2.1
 ; SEQ ID NO 103
 ; LENGTH: 591
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-688-188B-103

Query Match 100.0%; Score 3090; DB 2; Length 591;
 Best Local Similarity 100.0%; Pred. No. 6.6e-191;
 Matches 591; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MFGKRRKVEISAPSNFHRVHTGFDQHEQKFTGLPRQWQSLIESARRPKPLVDPAKIT 60
 Db 1 MFGKRRKVEISAPSNFHRVHTGFDQHEQKFTGLPRQWQSLIESARRPKPLVDPAKIT 60
 QY 61 SIQCAPKTIVRGSKAKDGLTLLDDEFNNVTSRNSLRDSDPPPARAQENGMPEE 120
 Db 61 SIQCAPKTIVRGSKAKDGLTLLDDEFNNVTSRNSLRDSDPPPARAQENGMPEE 120
 QY 121 PATTARGGPKAGSRGRFAGHSEAGGSGDRRRRAGPEKPKSSREGSGGPQESSDKRPL 180
 Db 121 PATTARGGPKAGSRGRFAGHSEAGGSGDRRRRAGPEKPKSSREGSGGPQESSDKRPL 180

Db 121 PATTARGGPKAGSRGRFAGHSEAGGSGDRRRRAGPEKPKSSREGSGGPQESSDKRPL 180
 QY 181 SGPDVGTTPQAGLASGAKLAAGRPNFTYPRADTDHPSRGAQCEPHDVAPNGPSAGGLAIP 240
 Db 181 SGPDVGTTPQAGLASGAKLAAGRPNFTYPRADTDHPSRGAQCEPHDVAPNGPSAGGLAIP 240
 QY 241 QSSSSSSRPPTRARGAPSGVLGPHASEPQLAPACTPAAPVPGPPGPRSPQRPQVRS 300
 Db 241 QSSSSSSRPPTRARGAPSGVLGPHASEPQLAPACTPAAPVPGPPGPRSPQRPQVRS 300
 QY 301 HEQFRAALQLVVDGPDPSYLDNFIKIGEGSTGIVCIATVRSSGKLVAVKMDLRKQRR 360
 Db 301 HEQFRAALQLVVDGPDPSYLDNFIKIGEGSTGIVCIATVRSSGKLVAVKMDLRKQRR 360
 QY 361 ELLFNEVIMRDYQHENVVEMYSYLVGDELWVMEFLEGGALTDIVTHTRMNEEQIAAV 420
 Db 361 ELLFNEVIMRDYQHENVVEMYSYLVGDELWVMEFLEGGALTDIVTHTRMNEEQIAAV 420
 QY 421 CLAVLQALSVLHAQGVTHRDIKSDSILLTHDGRVKLSDFGFCQAQVSKVPRKSLVGTPT 480
 Db 421 CLAVLQALSVLHAQGVTHRDIKSDSILLTHDGRVKLSDFGFCQAQVSKVPRKSLVGTPT 480
 QY 481 WMAPELISRLPYGPEVDIWSLGIWIMVMDGEPYFNEPPLKAMKMIKIRDLNPPRLKNLHK 540
 Db 481 WMAPELISRLPYGPEVDIWSLGIWIMVMDGEPYFNEPPLKAMKMIKIRDLNPPRLKNLHK 540
 QY 541 VSPSLKGFLLRDLVRDPAQRATAAELLKHPFLAKAGPPASIVPLMRQNRTR 591
 Db 541 VSPSLKGFLLRDLVRDPAQRATAAELLKHPFLAKAGPPASIVPLMRQNRTR 591

RESULT 3
 US-09-718-032-2
 ; Sequence 2, Application US/09718032
 ; Patent No. 6667168
 ; GENERAL INFORMATION:
 ; APPLICANT: The Trustees of Columbia University
 ; APPLICANT: Minden, Audrey
 ; TITLE OF INVENTION: PAK4, A NOVEL GENE ENCODING A SERINE/THREONINE KINASE
 ; FILE REFERENCE: 575/55311-A-PCT-US
 ; CURRENT APPLICATION NUMBER: US/09/718,032
 ; CURRENT FILING DATE: 2000-11-21
 ; PRIOR APPLICATION NUMBER: PCT/US99/11341
 ; PRIOR FILING DATE: 1999-05-21
 ; PRIOR APPLICATION NUMBER: 09/082,737
 ; PRIOR FILING DATE: 1998-05-21
 ; NUMBER OF SEQ ID NOS: 17
 ; SOFTWARE: Patent in version 3.1
 ; SEQ ID NO 2
 ; LENGTH: 591
 ; TYPE: PRT
 ; ORGANISM: human
 US-09-718-032-2

Query Match 100.0%; Score 3090; DB 2; Length 591;
 Best Local Similarity 100.0%; Pred. No. 6.6e-191;
 Matches 591; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MFGKRRKVEISAPSNFHRVHTGFDQHEQKFTGLPRQWQSLIESARRPKPLVDPAKIT 60
 Db 1 MFGKRRKVEISAPSNFHRVHTGFDQHEQKFTGLPRQWQSLIESARRPKPLVDPAKIT 60
 QY 61 SIQCAPKTIVRGSKAKDGLTLLDDEFNNVTSRNSLRDSDPPPARAQENGMPEE 120
 Db 61 SIQCAPKTIVRGSKAKDGLTLLDDEFNNVTSRNSLRDSDPPPARAQENGMPEE 120
 QY 121 PATTARGGPKAGSRGRFAGHSEAGGSGDRRRRAGPEKPKSSREGSGGPQESSDKRPL 180
 Db 121 PATTARGGPKAGSRGRFAGHSEAGGSGDRRRRAGPEKPKSSREGSGGPQESSDKRPL 180
 QY 181 SGPDVGTTPQAGLASGAKLAAGRPNFTYPRADTDHPSRGAQCEPHDVAPNGPSAGGLAIP 240
 Db 181 SGPDVGTTPQAGLASGAKLAAGRPNFTYPRADTDHPSRGAQCEPHDVAPNGPSAGGLAIP 240

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration, Ltd.

OM protein - protein search, using sw model

Run on: March 2, 2006, 18:40:20 ; Search time 215.348 Seconds
(without alignments)
1146.689 Million cell updates/sec

Title: US-10-693-367-2

Perfect score: 3090

Sequence: 1 MFGKKRVEISAPSNFEHR.....LAKAGPPASIVPLMRQNRTR 591

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA Main:

- 1: /cgn2_6/prodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/prodata/1/pubpaa/US08_PUBCOMB.pep.*
- 3: /cgn2_6/prodata/1/pubpaa/US09_PUBCOMB.pep.*
- 4: /cgn2_6/prodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 5: /cgn2_6/prodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 6: /cgn2_6/prodata/1/pubpaa/US11_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|--------|-------------|--------|----|---------------------|
| 1 | 3090 | 100.0 | 591 | 3 | US-09-291-417-103 |
| 2 | 3090 | 100.0 | 591 | 4 | US-10-134-102-4 |
| 3 | 3090 | 100.0 | 591 | 4 | US-10-394-322A-48 |
| 4 | 3090 | 100.0 | 591 | 4 | US-10-693-367-2 |
| 5 | 3090 | 100.0 | 591 | 5 | US-10-725-329-103 |
| 6 | 3090 | 100.0 | 591 | 5 | US-10-893-025-5 |
| 7 | 3090 | 100.0 | 591 | 5 | US-10-725-121-103 |
| 8 | 3090 | 100.0 | 591 | 5 | US-10-509-599-18 |
| 9 | 3090 | 100.0 | 620 | 5 | US-10-450-763-49667 |
| 10 | 3075 | 99.5 | 588 | 4 | US-10-134-102-1 |
| 11 | 2860 | 92.6 | 593 | 3 | US-10-618-941-111 |
| 12 | 2073 | 67.1 | 398 | 3 | US-09-291-417-30 |
| 13 | 2073 | 67.1 | 398 | 5 | US-10-725-329-30 |
| 14 | 2073 | 67.1 | 398 | 5 | US-10-725-121-30 |
| 15 | 1748 | 56.6 | 719 | 4 | US-10-331-095-2 |
| 16 | 1731 | 56.0 | 719 | 4 | US-10-331-095-4 |
| 17 | 1731 | 56.0 | 719 | 4 | US-10-394-322A-49 |
| 18 | 1502 | 48.6 | 292 | 4 | US-10-406-676-5 |
| 19 | 1502 | 48.6 | 292 | 4 | US-10-406-676-6 |
| 20 | 1502 | 48.6 | 292 | 4 | US-10-406-676-8 |
| 21 | 1502 | 48.6 | 292 | 4 | US-10-406-676-9 |
| 22 | 1502 | 48.6 | 292 | 4 | US-10-406-676-10 |
| 23 | 1502 | 48.6 | 292 | 4 | US-10-406-676-11 |
| 24 | 1502 | 48.6 | 293 | 4 | US-10-406-676-4 |
| 25 | 1497.5 | 48.5 | 639 | 4 | US-10-134-102-6 |
| 26 | 1494.5 | 48.4 | 635 | 4 | US-10-134-102-2 |
| 27 | 1481 | 47.9 | 292 | 4 | US-10-406-676-7 |

| | | | | | | |
|----|--------|------|-----|---|-------------------|--------------------|
| 28 | 1481 | 47.9 | 292 | 4 | US-10-406-676-12 | Sequence 12, Appl |
| 29 | 1473.5 | 47.7 | 681 | 3 | US-09-291-417-29 | Sequence 29, Appl |
| 30 | 1473.5 | 47.7 | 681 | 5 | US-10-725-329-29 | Sequence 29, Appl |
| 31 | 1473.5 | 47.7 | 681 | 5 | US-10-737-450-108 | Sequence 108, Appl |
| 32 | 1473.5 | 47.7 | 681 | 5 | US-10-885-921-10 | Sequence 10, Appl |
| 33 | 1473.5 | 47.7 | 681 | 3 | US-10-725-121-29 | Sequence 29, Appl |
| 34 | 1470.5 | 47.6 | 681 | 3 | US-09-765-815-2 | Sequence 2, Appl |
| 35 | 1467 | 47.5 | 682 | 5 | US-10-840-512-209 | Sequence 209, Appl |
| 36 | 1421 | 46.0 | 292 | 4 | US-10-406-676-15 | Sequence 15, Appl |
| 37 | 1335.5 | 43.2 | 641 | 4 | US-10-311-034-15 | Sequence 15, Appl |
| 38 | 1303 | 42.2 | 290 | 4 | US-10-406-676-13 | Sequence 13, Appl |
| 39 | 1289 | 41.7 | 290 | 4 | US-10-406-676-14 | Sequence 14, Appl |
| 40 | 1289 | 41.7 | 290 | 4 | US-10-406-676-16 | Sequence 16, Appl |
| 41 | 1289 | 41.7 | 290 | 4 | US-10-406-676-17 | Sequence 17, Appl |
| 42 | 1277 | 41.3 | 250 | 4 | US-10-693-367-3 | Sequence 3, Appl |
| 43 | 1213 | 39.3 | 290 | 4 | US-10-406-676-18 | Sequence 18, Appl |
| 44 | 1209 | 39.1 | 290 | 4 | US-10-406-676-19 | Sequence 19, Appl |
| 45 | 1209 | 39.1 | 290 | 4 | US-10-406-676-20 | Sequence 20, Appl |

ALIGNMENTS

RESULT 1

US-09-291-417-103

; Sequence 103, Application US/09291417A

; Publication No. US20030050230A1

; GENERAL INFORMATION:

; APPLICANT: PLOWMAN, GREGORY

; APPLICANT: MARTINEZ, RICARDO

; APPLICANT: WHYTE, DAVID

; TITLE OF INVENTION: STE20-RELATED PROTEIN KINASES

; FILE REFERENCE: 240/300

; CURRENT APPLICATION NUMBER: US/09/291,417A

; CURRENT FILING DATE: 1999-04-13

; EARLIER APPLICATION NUMBER: US 60/081,784

; EARLIER FILING DATE: 1998-04-14

; NUMBER OF SEQ ID NOS: 147

; SOFTWARE: FASTSEQ for Windows Version 3.0

; SEQ ID NO 103

; LENGTH: 591

; TYPE: PRF

; ORGANISM: Full Length Mammalian (Human) PAK5hu

US-09-291-417-103

| | | | | |
|-----------------------|-----------------|--|---------------|-------------------|
| Query Match | 100.0% | Score 3090; | DB 3; | Length 591; |
| Best Local Similarity | 100.0%; | Pred. No. 7.5e-161; | Mismatches 0; | Indels 0; Gaps 0; |
| Matches 591; | Conservative 0; | | | |
| QY | 1 | MFGKKRVEISAPSNFEHRVHTGPDQHEQKFTGLPRQWQSLIESARRPKPLVDPACIT | 60 | |
| Db | 1 | MFGKKRVEISAPSNFEHRVHTGPDQHEQKFTGLPRQWQSLIESARRPKPLVDPACIT | 60 | |
| QY | 61 | SIQCAPKTIVRGSKGAKDGALTLLDDEFNMVTRSNLRDSDPPPARAQENGMPEE | 120 | |
| Db | 61 | SIQCAPKTIVRGSKGAKDGALTLLDDEFNMVTRSNLRDSDPPPARAQENGMPEE | 120 | |
| QY | 121 | PATTARGGPKAGSRGRFAGHSEAGGSDRRRAGPEKPKSSRSGSGPQESSRDKRPL | 180 | |
| Db | 121 | PATTARGGPKAGSRGRFAGHSEAGGSDRRRAGPEKPKSSRSGSGPQESSRDKRPL | 180 | |
| QY | 181 | SGPDVGTTPQAGLASKAKLAAGRPNPTYPTRADTHPSRGAQCEPHDVPNGSAGGLAIP | 240 | |
| Db | 181 | SGPDVGTTPQAGLASKAKLAAGRPNPTYPTRADTHPSRGAQCEPHDVPNGSAGGLAIP | 240 | |
| QY | 241 | QSSSSSSPPTRARGAPSGVLGPHASBPQLAPPACTPAAPAVPGPPRSPRQRPVRS | 300 | |
| Db | 241 | QSSSSSSPPTRARGAPSGVLGPHASBPQLAPPACTPAAPAVPGPPRSPRQRPVRS | 300 | |
| QY | 301 | HEQFRAALQVVDGDPKSYLDNFIKIGEGSTGIVCIATVRSRSGKLVAKWMDLRKQRR | 360 | |
| Db | 301 | HEQFRAALQVVDGDPKSYLDNFIKIGEGSTGIVCIATVRSRSGKLVAKWMDLRKQRR | 360 | |

QY 361 ELLFNEVIMRDYQHENVVEMNSYLVGDELWVMEFLEGGALTDIVTHTRMNEEQIAAV 420
Db 361 ELLFNEVIMRDYQHENVVEMNSYLVGDELWVMEFLEGGALTDIVTHTRMNEEQIAAV 420
QY 421 CLAVLQALSVLHAQGVITHRDIKSDSILLTHDGRVKLSDFGCAQVSKVPRKSLVGTPT 480
Db 421 CLAVLQALSVLHAQGVITHRDIKSDSILLTHDGRVKLSDFGCAQVSKVPRKSLVGTPT 480
QY 481 WMAPELISRLPYGPEVDIWSLGIWVEMVDEGPPYFNEPPLKAMKMIKMDLPPRLKNLHK 540
Db 481 WMAPELISRLPYGPEVDIWSLGIWVEMVDEGPPYFNEPPLKAMKMIKMDLPPRLKNLHK 540
QY 541 VSPSLKGFLLVDRDPAQRATAAELLKHPFLAKAGPPASIVPLMRQNRTR 591
Db 541 VSPSLKGFLLVDRDPAQRATAAELLKHPFLAKAGPPASIVPLMRQNRTR 591

RESULT 2
US-10-134-102-4
; Sequence 4, Application US/10134102
; Publication No. US20030186254A1
; GENERAL INFORMATION:
; APPLICANT: Moritz, Michael B.
; APPLICANT: Moritz, Albrecht
; APPLICANT: Comb, Michael J.
; TITLE OF INVENTION: Regulation of HIV-Tat and Nef by the Pak4 kinase and its
; binding partners and methods of identifying modulators thereof.
; FILE REFERENCE: CST-176 CIP
; CURRENT APPLICATION NUMBER: US/10/134,102
; PRIOR FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 09/750,457
; PRIOR FILING DATE: 2000-12-28
; PRIOR APPLICATION NUMBER: 60/173,939
; PRIOR FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-134-102-4

Query Match 100.0%; Score 3090; DB 4; Length 591;
Best Local Similarity 100.0%; Pred. No. 7.5e-161;
Matches 591; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MFGKKRKEVEISAPSNFHRVHTGFDQHEQKFTGLPROMQSLIESARRPKPLVDPAIT 60
Db 1 MFGKKRKEVEISAPSNFHRVHTGFDQHEQKFTGLPROMQSLIESARRPKPLVDPAIT 60
QY 61 SIQCAPKTIIVRGSKGAKDGALTLLDEFENMVSRTSNLSLRDSDPPPPARARQENGMPPEE 120
Db 61 SIQCAPKTIIVRGSKGAKDGALTLLDEFENMVSRTSNLSLRDSDPPPPARARQENGMPPEE 120
QY 121 PATTARGGPKAGSRGAFAGHSEAGGSGDRRRRAGPEKPKSSREGSGGPOESSRDKRPL 180
Db 121 PATTARGGPKAGSRGAFAGHSEAGGSGDRRRRAGPEKPKSSREGSGGPOESSRDKRPL 180
QY 181 SGPDVGTTPQAGLAGAKLAAGRPFNTYPRADTDHPSRGAQGEHDVAPNGPSAGGLAIP 240
Db 181 SGPDVGTTPQAGLAGAKLAAGRPFNTYPRADTDHPSRGAQGEHDVAPNGPSAGGLAIP 240
QY 241 QSSSSSSRPPTTRARGAPSPVGLPHASEPQLAPACTPAAPAVPGPPGPRSPQRPQVRS 300
Db 241 QSSSSSSRPPTTRARGAPSPVGLPHASEPQLAPACTPAAPAVPGPPGPRSPQRPQVRS 300
QY 301 HEQFRAALQVLVDPGDRSYLDNFKIGEGSTGIVCIATVRSRSGKLVAVKMDLKKQRR 360
Db 301 HEQFRAALQVLVDPGDRSYLDNFKIGEGSTGIVCIATVRSRSGKLVAVKMDLKKQRR 360
QY 361 ELLFNEVIMRDYQHENVVEMNSYLVGDELWVMEFLEGGALTDIVTHTRMNEEQIAAV 420
Db 361 ELLFNEVIMRDYQHENVVEMNSYLVGDELWVMEFLEGGALTDIVTHTRMNEEQIAAV 420

QY 421 CLAVLQALSVLHAQGVITHRDIKSDSILLTHDGRVKLSDFGCAQVSKVPRKSLVGTPT 480
Db 421 CLAVLQALSVLHAQGVITHRDIKSDSILLTHDGRVKLSDFGCAQVSKVPRKSLVGTPT 480
QY 481 WMAPELISRLPYGPEVDIWSLGIWVEMVDEGPPYFNEPPLKAMKMIKMDLPPRLKNLHK 540
Db 481 WMAPELISRLPYGPEVDIWSLGIWVEMVDEGPPYFNEPPLKAMKMIKMDLPPRLKNLHK 540
QY 541 VSPSLKGFLLVDRDPAQRATAAELLKHPFLAKAGPPASIVPLMRQNRTR 591
Db 541 VSPSLKGFLLVDRDPAQRATAAELLKHPFLAKAGPPASIVPLMRQNRTR 591

RESULT 3
US-10-394-322A-48
; Sequence 48, Application US/10394322A
; Publication No. US20030232391A1
; GENERAL INFORMATION:
; APPLICANT: SENEZIS PHARMACEUTICALS, INC.
; APPLICANT: Prescott, John C.
; TITLE OF INVENTION: IDENTIFICATION OF KINASE INHIBITORS
; FILE REFERENCE: 39750-0006 US
; CURRENT APPLICATION NUMBER: US/10/394,322A
; CURRENT FILING DATE: 2003-03-20
; PRIOR APPLICATION NUMBER: US 60/366,892
; PRIOR FILING DATE: 2002-03-21
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 48
; LENGTH: 591
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-394-322A-48

Query Match 100.0%; Score 3090; DB 4; Length 591;
Best Local Similarity 100.0%; Pred. No. 7.5e-161;
Matches 591; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MFGKKRKEVEISAPSNFHRVHTGFDQHEQKFTGLPROMQSLIESARRPKPLVDPAIT 60
Db 1 MFGKKRKEVEISAPSNFHRVHTGFDQHEQKFTGLPROMQSLIESARRPKPLVDPAIT 60
QY 61 SIQCAPKTIIVRGSKGAKDGALTLLDEFENMVSRTSNLSLRDSDPPPPARARQENGMPPEE 120
Db 61 SIQCAPKTIIVRGSKGAKDGALTLLDEFENMVSRTSNLSLRDSDPPPPARARQENGMPPEE 120
QY 121 PATTARGGPKAGSRGAFAGHSEAGGSGDRRRRAGPEKPKSSREGSGGPOESSRDKRPL 180
Db 121 PATTARGGPKAGSRGAFAGHSEAGGSGDRRRRAGPEKPKSSREGSGGPOESSRDKRPL 180
QY 181 SGPDVGTTPQAGLAGAKLAAGRPFNTYPRADTDHPSRGAQGEHDVAPNGPSAGGLAIP 240
Db 181 SGPDVGTTPQAGLAGAKLAAGRPFNTYPRADTDHPSRGAQGEHDVAPNGPSAGGLAIP 240
QY 241 QSSSSSSRPPTTRARGAPSPVGLPHASEPQLAPACTPAAPAVPGPPGPRSPQRPQVRS 300
Db 241 QSSSSSSRPPTTRARGAPSPVGLPHASEPQLAPACTPAAPAVPGPPGPRSPQRPQVRS 300
QY 301 HEQFRAALQVLVDPGDRSYLDNFKIGEGSTGIVCIATVRSRSGKLVAVKMDLKKQRR 360
Db 301 HEQFRAALQVLVDPGDRSYLDNFKIGEGSTGIVCIATVRSRSGKLVAVKMDLKKQRR 360
QY 361 ELLFNEVIMRDYQHENVVEMNSYLVGDELWVMEFLEGGALTDIVTHTRMNEEQIAAV 420
Db 361 ELLFNEVIMRDYQHENVVEMNSYLVGDELWVMEFLEGGALTDIVTHTRMNEEQIAAV 420
QY 421 CLAVLQALSVLHAQGVITHRDIKSDSILLTHDGRVKLSDFGCAQVSKVPRKSLVGTPT 480
Db 421 CLAVLQALSVLHAQGVITHRDIKSDSILLTHDGRVKLSDFGCAQVSKVPRKSLVGTPT 480
QY 481 WMAPELISRLPYGPEVDIWSLGIWVEMVDEGPPYFNEPPLKAMKMIKMDLPPRLKNLHK 540
Db 481 WMAPELISRLPYGPEVDIWSLGIWVEMVDEGPPYFNEPPLKAMKMIKMDLPPRLKNLHK 540

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 2, 2006, 18:40:55 ; Search time 25.1078 Seconds
(without alignments)
470.772 Million cell updates/sec

Title: US-10-693-367-2
Perfect score: 3090
Sequence: 1 MFGKKRVEISAPSNFHR.....LAKGPPASIVPLMRQNR 591

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 135339 seqs, 20000136 residues

Total number of hits satisfying chosen parameters: 135339

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA New.*
1: /cgn2_6/prodata/2/pubpaa/US08_NEW_PUB.pep.*
2: /cgn2_6/prodata/2/pubpaa/US06_NEW_PUB.pep.*
3: /cgn2_6/prodata/2/pubpaa/US07_NEW_PUB.pep.*
4: /cgn2_6/prodata/2/pubpaa/PCT_NEW_PUB.pep.*
5: /cgn2_6/prodata/2/pubpaa/US05_NEW_PUB.pep.*
6: /cgn2_6/prodata/2/pubpaa/US10_NEW_PUB.pep.*
7: /cgn2_6/prodata/2/pubpaa/US11_NEW_PUB.pep.*
8: /cgn2_6/prodata/2/pubpaa/US60_NEW_PUB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|---------------------|--------------------|
| 1 | 3090 | 100.0 | 591 | US-10-770-726-71 | Sequence 71, Appl |
| 2 | 904 | 29.3 | 524 | US-10-878-556A-131 | Sequence 131, Appl |
| 3 | 893.5 | 28.9 | 545 | US-11-134-563-16 | Sequence 16, Appl |
| 4 | 846.5 | 27.4 | 1230 | US-11-087-099-1702 | Sequence 1702, Ap |
| 5 | 846.5 | 27.4 | 1230 | US-11-087-099-8922 | Sequence 8922, Ap |
| 6 | 845.5 | 27.4 | 642 | US-11-087-099-7738 | Sequence 7738, Ap |
| 7 | 837.5 | 27.1 | 644 | US-11-087-099-11826 | Sequence 11826, A |
| 8 | 549 | 17.8 | 855 | US-11-087-099-10166 | Sequence 10166, A |
| 9 | 521.5 | 16.9 | 825 | US-11-087-099-5581 | Sequence 5581, Ap |
| 10 | 516.5 | 16.7 | 825 | US-11-087-099-4218 | Sequence 4218, Ap |
| 11 | 503.5 | 16.3 | 487 | US-11-187-410-4 | Sequence 4, Appli |
| 12 | 495.5 | 16.0 | 547 | US-10-770-726-87 | Sequence 87, Appl |
| 13 | 487.5 | 15.8 | 1360 | US-11-241-056-14 | Sequence 14, Appl |
| 14 | 486 | 15.7 | 256 | US-10-877-346-72 | Sequence 72, Appl |
| 15 | 486 | 15.7 | 256 | US-11-113-424-183 | Sequence 183, Appl |
| 16 | 486 | 15.7 | 431 | US-11-152-366-29 | Sequence 29, Appl |
| 17 | 484.5 | 15.7 | 443 | US-11-152-366-30 | Sequence 30, Appl |
| 18 | 484 | 15.7 | 588 | US-10-821-234-1137 | Sequence 1137, Ap |
| 19 | 470.5 | 15.2 | 676 | US-11-087-099-8603 | Sequence 8603, Ap |
| 20 | 458 | 14.8 | 898 | US-11-099-691-7 | Sequence 7, Appli |
| 21 | 458 | 14.8 | 898 | US-11-124-367A-330 | Sequence 330, Appl |
| 22 | 458 | 14.8 | 898 | US-11-124-367A-333 | Sequence 333, Appl |
| 23 | 451.5 | 14.6 | 275 | US-11-241-056-8 | Sequence 8, Appli |
| 24 | 434 | 14.0 | 1236 | US-11-115-086-4 | Sequence 4, Appli |
| 25 | 431 | 13.9 | 266 | US-11-241-056-13 | Sequence 13, Appl |

| | | | | | | |
|----|-------|------|------|---|---------------------|--------------------|
| 26 | 407 | 13.2 | 256 | 6 | US-10-877-346-74 | Sequence 74, Appl |
| 27 | 398 | 12.9 | 1338 | 7 | US-11-087-099-11043 | Sequence 11043, A |
| 28 | 380 | 12.3 | 876 | 7 | US-11-087-099-10193 | Sequence 10193, A |
| 29 | 374 | 12.1 | 1133 | 6 | US-10-821-234-1219 | Sequence 1219, Ap |
| 30 | 370.5 | 12.0 | 406 | 7 | US-11-087-099-1555 | Sequence 1555, Ap |
| 31 | 369 | 11.9 | 238 | 7 | US-11-115-086-2 | Sequence 2, Appli |
| 32 | 368 | 11.9 | 404 | 7 | US-11-087-099-2344 | Sequence 2344, Ap |
| 33 | 368 | 11.9 | 424 | 7 | US-11-087-099-5546 | Sequence 5546, Ap |
| 34 | 368 | 11.9 | 425 | 7 | US-11-087-099-1247 | Sequence 1247, Ap |
| 35 | 367.5 | 11.9 | 406 | 7 | US-11-087-099-4651 | Sequence 4651, Ap |
| 36 | 367 | 11.9 | 349 | 7 | US-11-087-099-7872 | Sequence 7872, Ap |
| 37 | 362.5 | 11.7 | 231 | 7 | US-11-151-601-5 | Sequence 5, Appli |
| 38 | 362.5 | 11.7 | 231 | 7 | US-11-151-601-12 | Sequence 12, Appli |
| 39 | 362 | 11.7 | 1558 | 6 | US-10-329-258-14 | Sequence 14, Appli |
| 40 | 361.5 | 11.7 | 277 | 7 | US-11-151-601-4 | Sequence 4, Appli |
| 41 | 361.5 | 11.7 | 605 | 7 | US-11-087-099-3240 | Sequence 3240, Ap |
| 42 | 359 | 11.6 | 1832 | 7 | US-11-087-099-7762 | Sequence 7762, Ap |
| 43 | 358 | 11.6 | 278 | 6 | US-10-055-877-149 | Sequence 149, App |
| 44 | 358 | 11.6 | 278 | 7 | US-11-103-065-4 | Sequence 4, Appli |
| 45 | 358 | 11.6 | 278 | 7 | US-11-151-601-10 | Sequence 10, Appli |

ALIGNMENTS

RESULT 1
US-10-770-726-71
; Sequence 71, Application US/10770726
; Publication No. US20050266409A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Brown, Eugene
; APPLICANT: Liu, Wei
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSING, PREVENTING, AND TREATING
; TITLE OF INVENTION: CANCERS
; FILE REFERENCE: AM101079 (031896-010000)
; CURRENT APPLICATION NUMBER: US/10/770,726
; CURRENT FILING DATE: 2004-02-04
; NUMBER OF SEQ ID NOS: 48640
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 71
; LENGTH: 591
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-770-726-71

| | | | | |
|-----------------------|--------------|---|---------------|-------------------|
| Query Match | 100.0%; | Score 3090; | DB 6; | Length 591; |
| Best Local Similarity | 100.0%; | Pred. No. 2.2e-175; | Mismatches 0; | Indels 0; Gaps 0; |
| Matches 591; | Conservative | 0; | | |
| Qy | 1 | MFGKKRVEISAPSNFHRVHTGFDQHEOKFTGLPROMQSLIESARRPKPLVDPACIT | 60 | |
| Db | 1 | MFGKKRVEISAPSNFHRVHTGFDQHEOKFTGLPROMQSLIESARRPKPLVDPACIT | 60 | |
| Qy | 61 | SIQCAPKTIVRGSKGAKDGLTLLDEFNMSVTRSLRDSPPPPARARQNGMPBE | 120 | |
| Db | 61 | SIQCAPKTIVRGSKGAKDGLTLLDEFNMSVTRSLRDSPPPPARARQNGMPBE | 120 | |
| Qy | 121 | PATTARGPGKAGSRFAGHSEAGGSGDRRRRAGPEKPKSSRGSGGPPQSSSRKRPL | 180 | |
| Db | 121 | PATTARGPGKAGSRFAGHSEAGGSGDRRRRAGPEKPKSSRGSGGPPQSSSRKRPL | 180 | |
| Qy | 181 | SGPDVGTTPQAGLASKAKLAGRPENTYPRADTDHPSGACGEPHDVAPNGSAGGLATP | 240 | |
| Db | 181 | SGPDVGTTPQAGLASKAKLAGRPENTYPRADTDHPSGACGEPHDVAPNGSAGGLATP | 240 | |
| Qy | 241 | QSSSSSSRPPTARCAPSPGVLPASBPQLAPACTTAAAPVPGPPGPRSPQRQVRS | 300 | |
| Db | 241 | QSSSSSSRPPTARCAPSPGVLPASBPQLAPACTTAAAPVPGPPGPRSPQRQVRS | 300 | |
| Qy | 301 | HEQFRAALQLVVDPGDPRSYLDNFIKIGEGSTGIVCIATVRSSGKLAVKQMDLRQQR | 360 | |
| Db | 301 | HEQFRAALQLVVDPGDPRSYLDNFIKIGEGSTGIVCIATVRSSGKLAVKQMDLRQQR | 360 | |

QY 361 ELLFNEVIMRDYOHENVMYNSYLVGDELWVMEFELEGGALTDIVTHTRNNEQIAAV 420
Db 361 ELLFNEVIMRDYOHENVMYNSYLVGDELWVMEFELEGGALTDIVTHTRNNEQIAAV 420
QY 421 CLAVLQALSVLHAQGVHHRDIKSDSILLTHDGRVKLSDFGCAQVSKEVPRKSLVGTPY 480
Db 421 CLAVLQALSVLHAQGVHHRDIKSDSILLTHDGRVKLSDFGCAQVSKEVPRKSLVGTPY 480
QY 481 WMAPELISRLPYGPEVDIWSLIGIMVIMVGDGPPYFNBPPLKAMKMIKRDNLPPRLKLNHK 540
Db 481 WMAPELISRLPYGPEVDIWSLIGIMVIMVGDGPPYFNBPPLKAMKMIKRDNLPPRLKLNHK 540
QY 541 VSPSLKGFLLDRLLVDRDPAQRATAAELLKHPFLAKAGPPASIVPLMRQNRTR 591
Db 541 VSPSLKGFLLDRLLVDRDPAQRATAAELLKHPFLAKAGPPASIVPLMRQNRTR 591

RESULT 2

US-10-878-556A-131
; Sequence 131, Application US/10878556A
; Publication No. US20050266399A1
; GENERAL INFORMATION:
; APPLICANT: Hoffmann La-Roche Inc.
; TITLE OF INVENTION: HCV regulated protein expression
; FILE REFERENCE: 21762
; CURRENT APPLICATION NUMBER: US/10/878,556A
; CURRENT FILING DATE: 2004-06-28
; NUMBER OF SEQ ID NOS: 199
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 131
; LENGTH: 524
; TYPE: PRT
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: sw_hum/pak2_human
; DATABASE ENTRY DATE: 1997-11-01
US-10-878-556A-131

Query Match 29.3%; Score 904; DB 6; Length 524;
Best Local Similarity 35.0%; Pred. No. 1.6e-46;

Matches 211; Conservative 94; Mismatches 136; Indels 162; Gaps 13;

QY 4 KRKRVEISAPSNFHRVHTGFDQHEOKFTGLPRQWQSLIEESARRPKPLVDP 56
Db 67 KEKERPEISPPSDFEHTIHVGFDVATGFTGMPQEWARLLQTSNTKLEQKNPQAVLD- 125
QY 57 ACITSIQPGAPKTIVRGSKAGDKGALTLLLDDEFENMSTRNSLRD--SPPPRARQOE 114
Db 126 -----VLKPYD-----SNTVKQKYLSTFTPP-----EK 147
QY 115 NGMPE-EPATTARGPGKAGSRGRFAGHSEAGGSGDRRRRAGPKRPSREGSGPOES 173
Db 148 DGLPGSTPALNAGK-----TEA-----PAVVTSEEDDDET 178
QY 174 SRDKRPLSGPVGTPQAGLAKAGRPNTYPRADTDHPSRGAOGEHVDVAPNGFS 233
Db 179 A-----PPVIAPRPDHTKS-----IVTRSVID-PVPAPVGDSDH----- 210
QY 234 AGGLAIPOSSSSSRPPTTRAGAPSPGVLGPHASEPQLAPPACTPAAPVAPGPPGRSPQ 293
Db 211 -----VDGAASKLDQK 221
QY 294 REPQVRVSHQFRAALQVLVDPDPRSYLDFNFKIGEGSTGIVCIATVRSSGKLVAKKMD 353
Db 222 KKKPKMTDEIMEKLETVISIGDPKKYTRYEKIQGASGTFTATDVALGDEVALKQIN 281
QY 354 LRQQRRELLFNEVIMRDYOHENVMYNSYLVGDELWVMEFELEGGALTDIVTHTRN 413
Db 282 LQKQPKKELIINEILLVMKELKNPNIWNFLSDYLVGDELWVMEYLAGGSLTDVVTETCMD 341
QY 414 EQQIAVCLAVLQALSVLHAQGVHHRDIKSDSILLTHDGRVKLSDFGCAQVSKEVPRK 473

Db 342 EAQIAAIVCRECLQALEFLHANQVIHRDIKSDNVLGMBSGVKLTDFGCAQITPEQSKRS 401
QY 474 SLVGTPTYMABELISRLPYGPEVDIWSLIGIMVIMVGDGPPYFNBPPLKAMKMIKRDNLPP 533
Db 402 TMVGTPYMAPEVTVRKAYGPKVDIWSLIGIMVIMVGDGPPYFNBPPLKAMKMIKRDNLPP 461
QY 534 RLKLNHKVSPSLKGFLLDRLLVDRDPAQRATAAELLKHPFLAKAGPPASIVPL- 586
Db 462 ELQNPKEKLSPIPRDFLNKCLNDVEKRGSAKELLQHPFLKAKPLSLTPLIMAAKEAMK 521
QY 587 QNR 589
Db 522 SNR 524

RESULT 3

US-11-134-563-16
; Sequence 16, Application US/11134563
; Publication No. US20050287569A1
; GENERAL INFORMATION:
; APPLICANT: Leong, John M.
; APPLICANT: Campellone, Kenneth G.
; TITLE OF INVENTION: ESPFNUCLEIC ACIDS AND PROTEINS AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: 07917-280001
; CURRENT APPLICATION NUMBER: US/11/134,563
; CURRENT FILING DATE: 2005-05-20
; PRIOR APPLICATION NUMBER: US 60/573,600
; PRIOR FILING DATE: 2004-05-20
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 545
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-134-563-16

Query Match 28.9%; Score 893.5; DB 7; Length 545;
Best Local Similarity 36.9%; Pred. No. 6.8e-46;

Matches 216; Conservative 82; Mismatches 166; Indels 121; Gaps 11;

QY 4 KRKRVEISAPSNFHRVHTGFDQHEOKFTGLPRQWQSLIEESARRPKPLVDPACITSIQ 63
Db 68 KEKERPEISLPSDFEHTIHVGFDVATGFTGMPQEWARLLQTSN----- 111
QY 64 PCAPKTIVRGSKAGDKGALTLLLDDEFENMSTRNSLRDSPPPPARARQENGMPEPAT 123
Db 112 -----ITKSEQKNPQAVLDVLEFYNSKKT-SNSQKYS----- 144
QY 124 TARGPGKAGSRGRFAGHSEAGGSGDRRRRAGPKRPSREGSGGPOESSRDKRPLSGP 183
Db 145 -----FTDKSAEDYNSNALNV-----KAVSETPAVPPVSEDED---DDD 181
QY 184 DVGTPQAGLAKAGRPNT---YPRADTDHPSRGAOGEHVDVAPNGFSAGGLAIP 240
Db 182 DDATPPP-----VIAPRPEHTKSVYTRSVI-----EPLVFTPT----- 214
QY 241 OSSSSSSSRPPTTRAGAPSPGVLGPHASEPQLAPPACTPAAPVAPGPPGRSPRSPQRVS 300
Db 215 RDVATSPISPTENTTTP-----PDALTENT-----EKQKKKPKMS 249
QY 301 HEQFRAALQVLVDPDPRSYLDFNFKIGEGSTGIVCIATVRSSGKLVAKKMDLRQQR 360
Db 250 DEEILEKLRISVSGDPKKYTRYEKIQGASGTFTATDVALGDEVALKQIN 309
QY 361 ELLENVIMRDYOHENVMYNSYLVGDELWVMEFELEGGALTDIVTHTRNNEQIAAV 420
Db 310 ELTINEILLVMKELKNPNIWNFLSDYLVGDELWVMEYLAGGSLTDVVTETCMDEQIAAV 369
QY 421 CLAVLQALSVLHAQGVHHRDIKSDSILLTHDGRVKLSDFGCAQVSKEVPRKSLVGTPY 480
Db 370 CRECLQALEFLHSNQHVIHRDIKSDNVLGMBSGVKLTDFGCAQITPEQSKRSTVGTPY 429